



SEQUENCE LISTING

<110> Puchter, Holger
gen, Christian

<120> Recombination systems and methods for eliminating nucleic acid sequences from genome of eukaryotic organisms

<130> 53262-20031.00/13173-00010-US

<140> US 10/750,891

<141> 2004-01-05

<150> PCT/EP02/07281

<151> 2002-07-02

<150> DE 101 31 786.7

<151> 2001-07-04

<160> 30

<170> PatentIn version 3.3

<210> 1

<211> 788

<212> DNA

<213> *Saccharomyces cerevisiae*

<220>

<221> CDS

<222> (62)..(766)

<223> open reading frame coding for I-SceI

<400> 1

```
ggatccagta ctgtacctag aatacaaaga agaggaagaa gaaacctcta cagaagaagt 60
g atg aaa aac atc aaa aaa aac cag gta atg aac ctg ggt ccg aac tct 109
  Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro Asn Ser
    1         5         10        15
aaa ctg ctg aaa gaa tac aaa tcc cag ctg atc gaa ctg aac atc gaa 157
Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn Ile Glu
    20        25        30
cag ttc gaa gca ggt atc ggt ctg atc ctg ggt gat gct tac atc cgt 205
Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr Ile Arg
    35        40        45
tct cgt gat gaa ggt aaa acc tac tgt atg cag ttc gag tgg aaa aac 253
Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp Lys Asn
    50        55        60
aaa gca tac atg gac cac gta tgt ctg ctg tac gat cag tgg gta ctg 301
Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp Val Leu
    65        70        75        80
tcc ccg ccg cac aaa aaa gaa cgt gtt aac cac ctg ggt aac ctg gta 349
Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn Leu Val
    85        90        95
atc acc tgg ggc gcc cag act ttc aaa cac caa gct ttc aac aaa ctg 397
Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn Lys Leu
    100       105       110
gct agc ctg ttc atc gtt aac aac aaa aaa acc atc ccg aac aac ctg 445
Ala Ser Leu Phe Ile Val Asn Asn Lys Lys Thr Ile Pro Asn Asn Leu
    115       120       125
```

gtt gaa aac tac ctg acc ccg atg tct ctg gca tac tgg ttc atg gat 493
 Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe Met Asp
 130 135 140
 gat ggt ggt aaa tgg gat tac aac aaa aac tct acc aac aaa tcg atc 541
 Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Thr Asn Lys Ser Ile
 145 150 155 160
 gta ctg aac acc cag tct ttc act ttc gaa gaa gta gaa tac ctg gtt 589
 Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Glu Tyr Leu Val
 165 170 175
 aag ggt ctg cgt aac aaa ttc caa ctg aac tgt tac cta aaa atc aac 637
 Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Leu Lys Ile Asn
 180 185 190
 aaa aac aaa ccg atc atc tac atc gat tct atg tct tac ctg atc ttc 685
 Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu Ile Phe
 195 200 205
 tac aac ctg atc aaa ccg tac ctg atc ccg cag atg atg tac aaa ctg 733
 Tyr Asn Leu Ile Lys Pro Tyr Leu Ile Pro Gln Met Met Tyr Lys Leu
 210 215 220
 ccg aac act atc tcc tcc gaa act ttc ctg aaa taataagtcg agtactggat 786
 Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys
 225 230 235
 cc 788

<210> 2

<211> 235

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 2

Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro Asn Ser
 1 5 10 15
 Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn Ile Glu
 20 25 30
 Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr Ile Arg
 35 40 45
 Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp Lys Asn
 50 55 60
 Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp Val Leu
 65 70 75 80
 Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn Leu Val
 85 90 95
 Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn Lys Leu
 100 105 110
 Ala Ser Leu Phe Ile Val Asn Asn Lys Lys Thr Ile Pro Asn Asn Leu
 115 120 125
 Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe Met Asp
 130 135 140
 Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Thr Asn Lys Ser Ile
 145 150 155 160
 Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Glu Tyr Leu Val
 165 170 175
 Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Leu Lys Ile Asn
 180 185 190
 Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu Ile Phe
 195 200 205
 Tyr Asn Leu Ile Lys Pro Tyr Leu Ile Pro Gln Met Met Tyr Lys Leu
 210 215 220
 Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys
 225 230 235

```
<220>
<221> CDS
<222> (54)..(737)
<223> open reading frame of I-ChuI with nuclear location
      signal
```

```
<220>
<221> misc_feature
<222> (54)..(83)
<223> coding for nuclear location signal
```

<400> 3																
ctc	gag	tacc	taga	ata	caa	aga	agag	ggaa	gaagaa	actc	tataga	agaa	gcc	atg	Met	56
															1	
ggt	cca	aag	aaa	aag	aga	aag	gtt	atc	atg	tca	tta	aca	caa	caa	caa	104
Gly	Pro	Lys	Lys	Lys	Arg	Lys	Val	Ile	Met	Ser	Leu	Thr	Gln	Gln	Gln	
			5				10				15					
aaa	gac	tta	att	ttc	gga	tct	cta	ctg	ggt	gat	gga	aat	tta	caa	act	152
Lys	Asp	Leu	Ile	Phe	Gly	Ser	Leu	Leu	Gly	Asp	Gly	Asn	Leu	Gln	Thr	
			20				25				30					
ggt	tca	gta	ggt	agg	act	tgg	cgc	tat	cga	gcg	ctc	cat	aaa	agt	gag	200
Gly	Ser	Val	Gly	Arg	Thr	Trp	Arg	Tyr	Arg	Ala	Leu	His	Lys	Ser	Glu	
			35				40				45					
cat	cag	aca	tac	tta	ttt	cat	aag	tat	gaa	atc	tta	aag	ccg	ctt	tgt	248
His	Gln	Thr	Tyr	Leu	Phe	His	Lys	Tyr	Glu	Ile	Leu	Lys	Pro	Leu	Cys	
			50				55				60					
ggc	gaa	aat	act	ctc	cca	aca	gaa	agt	ata	gtg	ttc	gac	gaa	aga	aca	296
Gly	Glu	Asn	Thr	Leu	Pro	Thr	Glu	Ser	Ile	Val	Phe	Asp	Glu	Arg	Thr	
			70				75				80					
aac	aag	gag	gtt	aaa	cgt	tgg	ttt	ttc	aac	aca	tta	acc	aat	cct	tcc	344
Asn	Lys	Glu	Val	Lys	Arg	Trp	Phe	Phe	Asn	Thr	Leu	Thr	Asn	Pro	Ser	
			85				90				95					
tta	aaa	ttc	ttc	gca	gac	atg	ttc	tac	aca	tat	gac	caa	aac	aca	caa	392
Leu	Lys	Phe	Phe	Ala	Asp	Met	Phe	Tyr	Thr	Tyr	Asp	Gln	Asn	Thr	Gln	
			100				105				110					
aaa	tgg	gtt	aaa	gat	gta	cct	gta	aag	gtt	caa	aca	ttc	tta	act	cct	440
Lys	Trp	Val	Lys	Asp	Val	Pro	Val	Lys	Val	Gln	Thr	Phe	Leu	Thr	Pro	
			115				120				125					
caa	gct	tta	gca	tac	ttt	tat	ata	gac	gat	gga	gcg	tta	aaa	tgg	ctt	488
Gln	Ala	Leu	Ala	Tyr	Phe	Tyr	Ile	Asp	Asp	Gly	Ala	Leu	Lys	Trp	Leu	
			130				135				140					
aat	aag	tct	aac	gct	atg	caa	att	tgt	act	gaa	agt	ttc	agt	caa	ggg	536
Asn	Lys	Ser	Asn	Ala	Met	Gln	Ile	Cys	Thr	Glu	Ser	Phe	Ser	Gln	Gly	
			150				155				160					
ggc	acg	att	cgg	atc	caa	aaa	gca	cta	aaa	acg	ctc	tat	aat	att	gat	584
Gly	Thr	Ile	Arg	Ile	Gln	Lys	Ala	Leu	Lys	Thr	Leu	Tyr	Asn	Ile	Asp	
			165				170				175					
aca	acg	ttg	aca	aaa	aaa	act	cta	caa	gac	ggc	aga	att	ggc	tat	cgt	632
Thr	Thr	Leu	Thr	Lys	Lys	Thr	Leu	Gln	Asp	Gly	Arg	Ile	Gly	Tyr	Arg	
			180				185				190					
ata	gct	att	cct	gaa	gcc	agt	agc	ggt	gct	ttt	cgt	gaa	gtc	att	aaa	680
Ile	Ala	Ile	Pro	Glu	Ala	Ser	Ser	Gly	Ala	Phe	Arg	Glu	Val	Ile	Lys	
			195				200				205					
cct	ttt	cta	gtt	gat	tgt	atg	aga	tac	aaa	gtt	tct	gat	ggc	aat	aaa	728

Pro Phe Leu Val Asp Cys Met Arg Tyr Lys Val Ser Asp Gly Asn Lys
 210 215 220 225
 ggc cac ctt tagctcgag
 Gly His Leu

746

<210> 4
 <211> 228
 <212> PRT
 <213> Chlamydomonas applanata

<400> 4
 Met Gly Pro Lys Lys Lys Arg Lys Val Ile Met Ser Leu Thr Gln Gln
 1 5 10 15
 Gln Lys Asp Leu Ile Phe Gly Ser Leu Leu Gly Asp Gly Asn Leu Gln
 20 25 30
 Thr Gly Ser Val Gly Arg Thr Trp Arg Tyr Arg Ala Leu His Lys Ser
 35 40 45
 Glu His Gln Thr Tyr Leu Phe His Lys Tyr Glu Ile Leu Lys Pro Leu
 50 55 60
 Cys Gly Glu Asn Thr Leu Pro Thr Glu Ser Ile Val Phe Asp Glu Arg
 65 70 75 80
 Thr Asn Lys Glu Val Lys Arg Trp Phe Phe Asn Thr Leu Thr Asn Pro
 85 90 95
 Ser Leu Lys Phe Phe Ala Asp Met Phe Tyr Thr Tyr Asp Gln Asn Thr
 100 105 110
 Gln Lys Trp Val Lys Asp Val Pro Val Lys Val Gln Thr Phe Leu Thr
 115 120 125
 Pro Gln Ala Leu Ala Tyr Phe Tyr Ile Asp Asp Gly Ala Leu Lys Trp
 130 135 140
 Leu Asn Lys Ser Asn Ala Met Gln Ile Cys Thr Glu Ser Phe Ser Gln
 145 150 155 160
 Gly Gly Thr Ile Arg Ile Gln Lys Ala Leu Lys Thr Leu Tyr Asn Ile
 165 170 175
 Asp Thr Thr Leu Thr Lys Lys Thr Leu Gln Asp Gly Arg Ile Gly Tyr
 180 185 190
 Arg Ile Ala Ile Pro Glu Ala Ser Ser Gly Ala Phe Arg Glu Val Ile
 195 200 205
 Lys Pro Phe Leu Val Asp Cys Met Arg Tyr Lys Val Ser Asp Gly Asn
 210 215 220
 Lys Gly His Leu
 225

<210> 5
 <211> 582
 <212> DNA
 <213> Chlamydomonas reinhardtii

<220>
 <221> CDS
 <222> (55)..(573)
 <223> openreading frame coding for I-CreI with nuclear
 location signal

<220>
 <221> misc_feature
 <222> (55)..(84)
 <223> coding for nuclear location signal

<400> 5

```

ctcgagtacc tagaatacaa agaagaggaa gagaaacctc taccagaaga agcc atg 57
Met
1
ggt cca aag aaa aag aga aag gtt atc atg aat aca aaa tat aat aaa 105
Gly Pro Lys Lys Lys Arg Lys Val Ile Met Asn Thr Lys Tyr Asn Lys
5 10 15
gag ttc tta ctc tac tta gca ggg ttt gta gac ggt gac ggt agc ata 153
Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser Ile
20 25 30
atc gct caa att aag cct aat cag tct tat aaa ttt aag cat cag cta 201
Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Gln Leu
35 40 45
tca ctc gcg ttc caa gtc acg caa aag aca cag aga cgt tgg ttt tta 249
Ser Leu Ala Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe Leu
50 55 60 65
gac aaa tta gtg gat gaa att ggg gtt ggt tat gta aga gat agg ggt 297
Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp Arg Gly
70 75 80
agc gtt tcg gat tat att cta agc gaa atc aag cct ttg cat aat ttt 345
Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile Lys Pro Leu His Asn Phe
85 90 95
tta aca caa cta caa cct ttt cta aaa cta aaa caa aaa caa gca aat 393
Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala Asn
100 105 110
tta gtt tta aaa att att gaa caa ctt ccg tca gca aaa gaa tcc ccg 441
Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser Pro
115 120 125
gac aaa ttc tta gaa gtt tgt aca tgg gtg gat caa att gca gct ctg 489
Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala Leu
130 135 140 145
aat gat tcg aag acg cgt aaa aca act tct gaa acc gtt cgt gct gtg 537
Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala Val
150 155 160
cta gac agt tta agt gaa aaa aag aaa tcg tcc ccg tagctcgag 582
Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro
165 170

```

<210> 6

<211> 173

<212> PRT

<213> Chlamydomonas reinhardtii

<400> 6

```

Met Gly Pro Lys Lys Lys Arg Lys Val Ile Met Asn Thr Lys Tyr Asn
1 5 10 15
Lys Glu Phe Leu Leu Tyr Leu Ala Gly Phe Val Asp Gly Asp Gly Ser
20 25 30
Ile Ile Ala Gln Ile Lys Pro Asn Gln Ser Tyr Lys Phe Lys His Gln
35 40 45
Leu Ser Leu Ala Phe Gln Val Thr Gln Lys Thr Gln Arg Arg Trp Phe
50 55 60
Leu Asp Lys Leu Val Asp Glu Ile Gly Val Gly Tyr Val Arg Asp Arg
65 70 75 80
Gly Ser Val Ser Asp Tyr Ile Leu Ser Glu Ile Lys Pro Leu His Asn
85 90 95
Phe Leu Thr Gln Leu Gln Pro Phe Leu Lys Leu Lys Gln Lys Gln Ala
100 105 110
Asn Leu Val Leu Lys Ile Ile Glu Gln Leu Pro Ser Ala Lys Glu Ser

```

115	120	125
Pro Asp Lys Phe Leu Glu Val Cys Thr Trp Val Asp Gln Ile Ala Ala		
130	135	140
Leu Asn Asp Ser Lys Thr Arg Lys Thr Thr Ser Glu Thr Val Arg Ala		
145	150	155
Val Leu Asp Ser Leu Ser Glu Lys Lys Lys Ser Ser Pro		160
165	170	

<210> 7
 <211> 546
 <212> DNA
 <213> Chlamydomonas segnis

<220>
 <221> CDS
 <222> (52)..(537)
 <223> open readings frame coding for I-CpaI with nuclear
 location signal

<220>
 <221> misc_feature
 <222> (52)..(81)
 <223> coding for nuclear location signal

<400> 7
 ctcgagtacc tagaatacaa gaagagaaga agaacctcta cagaagaagc c atg ggt 57
 Met Gly
 1
 cca aag aaa aag aga aag gtt atc atg gac att aat cct caa tgg att 105
 Pro Lys Lys Lys Arg Lys Val Ile Met Asp Ile Asn Pro Gln Trp Ile
 5 10 15
 aca ggt ttc gta gat ggg gaa ggt tgt ttt agt gta agt att ctt aga 153
 Thr Gly Phe Val Asp Gly Glu Gly Cys Phe Ser Val Ser Ile Leu Arg
 20 25 30
 aat aat tcg ttg cgc tat ggc cat cag ctt caa cca gaa ttc gta gtg 201
 Asn Asn Ser Leu Arg Tyr Gly His Gln Leu Gln Pro Glu Phe Val Val
 35 40 45 50
 acc caa cat aaa tta gat gca aat gtt tta tat gca tta aaa gac tac 249
 Thr Gln His Lys Leu Asp Ala Asn Val Leu Tyr Ala Leu Lys Asp Tyr
 55 60 65
 ttt aaa gtt gga tca gtc gtt gtg aat cat ggg gaa cgg ctt tgc tat 297
 Phe Lys Val Gly Ser Val Val Val Asn His Gly Glu Arg Leu Cys Tyr
 70 75 80
 aaa gtc aaa aat att gat cac ttt ata acc gtc att ata cca ttt ttc 345
 Lys Val Lys Asn Ile Asp His Phe Ile Thr Val Ile Ile Pro Phe Phe
 85 90 95
 gaa aaa cat gag cta aaa aca aaa aga aga att gaa ttt ctt cga ttt 393
 Glu Lys His Glu Leu Lys Thr Lys Arg Arg Ile Glu Phe Leu Arg Phe
 100 105 110
 cga aaa atc tgc ttg ctg tta aaa gca ggt aga cat tta gaa tcg cag 441
 Arg Lys Ile Cys Leu Leu Leu Lys Ala Gly Arg His Leu Glu Ser Gln
 115 120 125 130
 gaa gga ttc gag aaa gtg ttg gat tta gca aaa aaa ctc cgt atc aat 489
 Glu Gly Phe Glu Lys Val Leu Asp Leu Ala Lys Lys Leu Arg Ile Asn
 135 140 145
 gag aaa aac tac cag gaa tct atc aaa cgt ttt gaa gaa act ggc gag 537
 Glu Lys Asn Tyr Gln Glu Ser Ile Lys Arg Phe Glu Glu Thr Gly Glu
 150 155 160
 taactcgag 546

[illegible]

```
<210> 9
<211> 793
<212> DNA
<213> Chlamydomonas seignis
```

```
<220>
<221> CDS
<222> (53)..(784)
<223> open reading frame coding for I-CpaII with nuclear
location signal
```

```
<220>
<221> misc_feature
<222> (53)..(82)
<223> coding for nuclear location signal
```

<400> 9																	
ctcgagtacc tagaaacaaa gaagaggaag aagaaaactct acagaagaag cc atg ggt																	58
																	Met Gly
																	1
cca aag aaa aag aga aag gtt atc atg acc gat tct aaa tct aga aac	106																
Pro Lys Lys Lys Arg Lys Val Ile Met Thr Asp Ser Lys Ser Arg Asn																	
	5 10 15																
aac aat aat ttt tta agc aat aat ctt tta cct ttg acc gat gac gag	154																
Asn Asn Asn Phe Leu Ser Asn Asn Leu Leu Pro Leu Thr Asp Asp Glu																	
	20 25 30																
aag gct tta att gcg ggg aca ctt tta ggg gat gct cat att caa aag	202																
Lys Ala Leu Ile Ala Gly Thr Leu Leu Gly Asp Ala His Ile Gln Lys																	
	35 40 45 50																

```

cgt ggt gat agc tat agg cta aaa ata gct cat ggc ttg gat cat gaa 250
Arg Gly Asp Ser Tyr Arg Leu Lys Ile Ala His Gly Leu Asp His Glu
55 60 65
gag ctt gtc gtc tgg aag tat aac cgt tta atc agg ttg tgt caa aca 298
Glu Leu Val Val Trp Lys Tyr Asn Arg Leu Ile Arg Leu Cys Gln Thr
70 75 80
aca caa ccc cca agg gtg gaa acc tac tca aca aag tta aag tct ggc 346
Thr Gln Pro Pro Arg Val Glu Thr Tyr Ser Thr Lys Leu Lys Ser Gly
85 90 95
gta ttg cct caa ggg gtt gtt ttc tat acc tcg tcc gga aag tat tta 394
Val Leu Pro Gln Gly Val Val Phe Tyr Thr Ser Ser Gly Lys Tyr Leu
100 105 110
aaa gag act tat gac ctt ttt tat aaa caa act gca gac ggt cgg agg 442
Lys Glu Thr Tyr Asp Leu Phe Tyr Lys Gln Thr Ala Asp Gly Arg Arg
115 120 125 130
gta aaa aca ata aca cag gag ttg atc gac agt tta ccc aag cat cca 490
Val Lys Thr Ile Thr Gln Glu Leu Ile Asp Ser Leu Pro Lys His Pro
135 140 145
ttg gtc tta gca gcc ttt ttt atg gac gat ggt agt gtt cgg tcc gac 538
Leu Val Leu Ala Ala Phe Phe Met Asp Asp Gly Ser Val Arg Ser Asp
150 155 160
tgt tat tca gga aag att gca acg cca ggg ttt gct ggt aaa gaa gaa 586
Cys Tyr Ser Gly Lys Ile Ala Thr Pro Gly Phe Ala Gly Lys Glu Glu
165 170 175
agc cag ttg ttg tgt aac tat cta cac agt tgg gat gtt caa gca aac 634
Ser Gln Leu Leu Cys Asn Tyr Leu His Ser Trp Asp Val Gln Ala Asn
180 185 190
gta gtt gct cat aaa aaa gca aac aat cag tat tac att ggg ctc cca 682
Val Val Ala His Lys Lys Ala Asn Asn Gln Tyr Tyr Ile Gly Leu Pro
195 200 205 210
gca aaa aca ttt ggt cgc ttt att aac att att gaa ccc tac gtt aga 730
Ala Lys Thr Phe Gly Arg Phe Ile Asn Ile Ile Glu Pro Tyr Val Arg
215 220 225
gaa gtt cct gct tta tgt tat aaa tta aac gaa tca aga aaa ccc cgt 778
Glu Val Pro Ala Leu Cys Tyr Lys Leu Asn Glu Ser Arg Lys Pro Arg
230 235 240
aac gac tgactcgag 793
Asn Asp

```

<210> 10

<211> 244

<212> PRT

<213> Chlamydomonas segnis

<400> 10

```

Met Gly Pro Lys Lys Lys Arg Lys Val Ile Met Thr Asp Ser Lys Ser
1 5 10 15
Arg Asn Asn Asn Asn Phe Leu Ser Asn Asn Leu Leu Pro Leu Thr Asp
20 25 30
Asp Glu Lys Ala Leu Ile Ala Gly Thr Leu Leu Gly Asp Ala His Ile
35 40 45
Gln Lys Arg Gly Asp Ser Tyr Arg Leu Lys Ile Ala His Gly Leu Asp
50 55 60
His Glu Glu Leu Val Val Trp Lys Tyr Asn Arg Leu Ile Arg Leu Cys
65 70 75 80
Gln Thr Thr Gln Pro Arg Val Glu Thr Tyr Ser Thr Lys Leu Lys
85 90 95
Ser Gly Val Leu Pro Gln Gly Val Val Phe Tyr Thr Ser Ser Gly Lys
100 105 110

```


Tyr Leu Lys Glu Thr Tyr Asp Leu Phe Tyr Lys Gln Thr Ala Asp Gly
 115 120 125
 Arg Arg Val Lys Thr Ile Thr Gln Glu Leu Ile Asp Ser Leu Pro Lys
 130 135 140
 His Pro Leu Val Leu Ala Ala Phe Phe Met Asp Asp Gly Ser Val Arg
 145 150 155 160
 Ser Asp Cys Tyr Ser Gly Lys Ile Ala Thr Pro Gly Phe Ala Gly Lys
 165 170 175
 Glu Glu Ser Gln Leu Leu Cys Asn Tyr Leu His Ser Trp Asp Val Gln
 180 185 190
 Ala Asn Val Val Ala His Lys Lys Ala Asn Asn Gln Tyr Tyr Ile Gly
 195 200 205
 Leu Pro Ala Lys Thr Phe Gly Arg Phe Ile Asn Ile Ile Glu Pro Tyr
 210 215 220
 Val Arg Glu Val Pro Ala Leu Cys Tyr Lys Leu Asn Glu Ser Arg Lys
 225 230 235 240
 Pro Arg Asn Asp

<210> 11
 <211> 35
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of the artificial sequence:
 oligonucleotide primer

<400> 11
 cggctcgagc tacggggacg atttcttttt ttac

35

<210> 12
 <211> 120
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of the artificial sequence:
 oligonucleotide primer

<400> 12
 cggctcgagt acctagaata caaagaagag gaagaagaaa cctctacaga agaagccatg 60
 ggtccaaaga aaaagagaaa gggtatcatg aatacaaaat ataataaaga gttcttactc 120

<210> 13
 <211> 116
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of the artificial sequence:
 oligonucleotide primer

<400> 13
 cggctcgagt acctagaata caaagaagag gaagaagaaa cctctacaga agaagccatg 60
 ggtccaaaga aaaagagaaa gggtatcatg gacattaatc ctcaatggat tacagg 116

<210> 14
 <211> 33
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of the artificial sequence:
 oligonucleotide primer

<400> 14
 cggctcgagt tactcgccag tttcttcaaa acg 33

<210> 15
 <211> 114
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of the artificial sequence:
 oligonucleotide primer

<400> 15
 cggctcgagt acctagaata caaagaagag gaagaagaaa cctctacaga agaagccatg 60
 ggtccaaaga aaaagagaaa gggtatcatg accgattcta aatctagaaa caac 114

<210> 16
 <211> 35
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of the artificial sequence:
 oligonucleotide primer

<400> 16
 cggctcgagc taaaggtggc ctttattgcc atcag 35

<210> 17
 <211> 114
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of the artificial sequence:
 oligonucleotide primer

<400> 17
 cggctcgagt acctagaata caaagaagag gaagaagaaa cctctacaga agaagccatg 60
 ggtccaaaga aaaagagaaa gggtatcatg tcattaacac aacaacaaaa agac 114

<210> 18
 <211> 35
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Description of the artificial sequence:
oligonucleotide primer

<400> 18

cggctcgagc taaaggtggc ctttattgcc atcag

35

<210> 19

<211> 66

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of the artificial sequence:
oligonucleotide primer

<400> 19

cggctctaga gcggccgcct agggataaca gggtaataga atcccacaaa aatctgagct 60
taacag 66

<210> 20

<211> 65

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of the artificial sequence:
oligonucleotide primer

<400> 20

cggctctaga ctattaccct gttatcccta ggcccgatct agtaacatag atgacaccgc 60
gcgcg 65

<210> 21

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of the artificial sequence:
oligonucleotide primer

<400> 21

cggaagcttc gtcaccaatc ccaattcgat ctac

34

<210> 22

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of the artificial sequence:
oligonucleotide primer

<400> 22

cggaagcttc cacttgcaaa gtcccgctag tgcc

34

<210> 23
 <211> 34
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of the artificial sequence:
 oligonucleotide primer

<400> 23
 cggaagcttc gtcaccaatc ccaattcgat ctac 34

<210> 24
 <211> 34
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of the artificial sequence:
 oligonucleotide primer

<400> 24
 cggaagcttc cacttgcaaa gtcccgctag tgcc 34

<210> 25
 <211> 62
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of the artificial sequence:
 oligonucleotide primer

<400> 25
 ctagtacaaa acgtcgtgag acattttaat ctgaaggttt ggcacctcga tgcgggtca 60
 tc 62

<210> 26
 <211> 62
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of the artificial sequence:
 oligonucleotide primer

<400> 26
 ctaggatgag ccgtcatcga ggtgccaaac cttcagatta aaatgtctca cgacgttttg 60
 ta 62

<210> 27
 <211> 75
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Description of the artificial sequence:
oligonucleotide primer

<400> 27

ctagtccgaa aacgccgtga gacatatggg ttacgatacct aaggtagcga aattcacccg 60
gtaactctgt gccag 75

<210> 28

<211> 75

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of the artificial sequence:
oligonucleotide primer

<400> 28

ctagctggca cagagttacc gggatgaattt cgctaccta ggatcgtaac caatatgtct 60
cacggcggtt tcgga 75

<210> 29

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of the artificial sequence: nuclear
location sequence

<400> 29

Pro Lys Thr Lys Arg Lys Val
1 5

<210> 30

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of the artificial sequence: nuclear
location sequence

<400> 30

Pro Lys Lys Lys Arg Lys Val
1 5